Virtual Hot Spare Disk Units

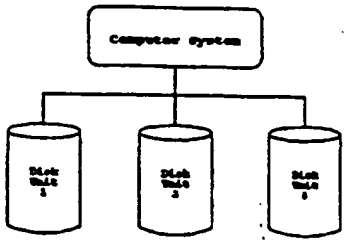

Fig. 1a

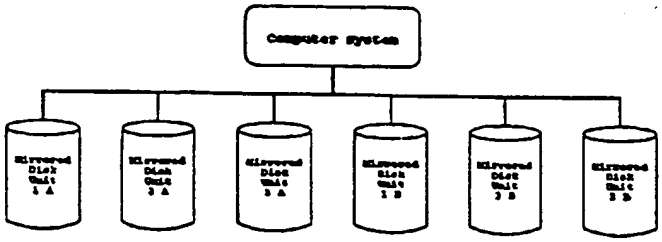

Fig. 1b

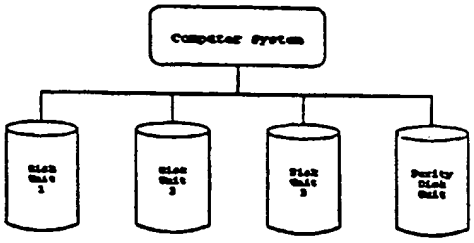

Fig. 1c

Fig. 1. Computer Systems With No Spare Disk Units.

When a disk unit attached to a computer system (Fig. 1a) fails, the data stored on that disk unit may be lost and/or availability of the computer system to do useful work may be temporarily lost. If the disk unit is protected using mirroring (Fig. 1b) or RAID (Fig. 1c) techniques, the protection becomes exposed when a disk unit fails, and then data and/or system availability may be lost if a second disk unit fails prior to the repair or replacement of the first failed unit. The potential for data or system availability loss may be reduced if a spare disk unit is attached to the computer system. A method is described for providing the spare disk unit capability without actually requiring that a specifically designated spare disk unit be attached to the computer system.

The exposure to loss of data and/or system availability due to disk unit failure may be reduced through the use of "hot spare disk units." The hot spare disk units approach typically involves attaching extra, unused, inactive disk units to the computer system (Figs. 2a, 2b, 2c).

Then, when one of the active disk units signals that it is about to fail (or, in the mirrored or RAID case, actually does fail), the data stored on that unit is moved onto the spare unit, thus avoiding the loss of data and/or loss of system availability (or, in the mirrored or RAID case, minimizing the duration of exposure to such loss due to the failure of a second unit). However, if multiple different disk unit types or sizes are attached to the computer system, then this approach to hot spare disk units will require attaching multiple spare disk units -- one of each type or size. Also, if one desires to protect against the failure of multiple active disk units, then multiple spare disk units must be attached. Finally, since the attached spare disk units are inactive, they do not contribute the usual improvement to system performance expected when additional disk units are attached.

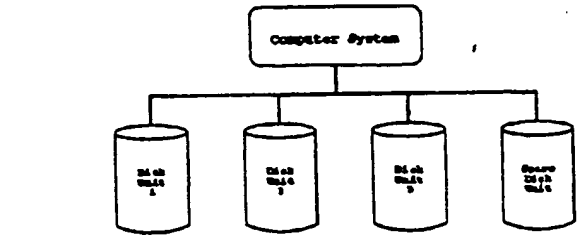

Fig. 2a

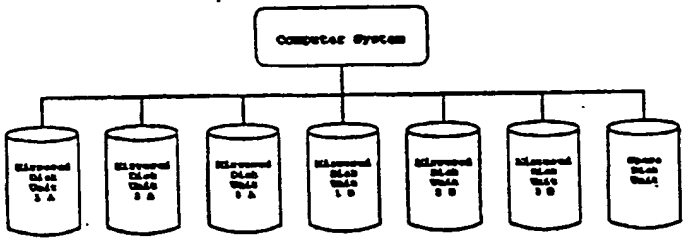

Fig. 2b

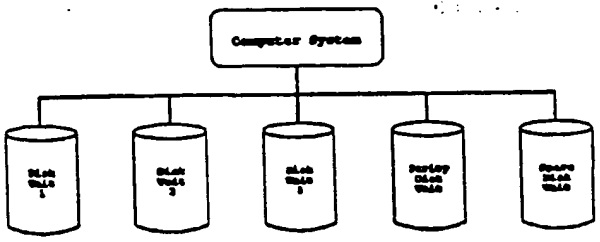

Fig. 2c

Fig. 2. Computer Systems With Hot Spare Disk Units.

The method of reducing exposure to loss of disk unit data and/or system availability referred to as "virtual hot spare disk units" only requires that there be adequate unused disk storage capacity on the active disk units attached to the computer system to contain the data from any single disk unit attached (Figs. 3a, 3b, 3c). If the disk units on the system are grouped into logical pools, then there must be adequate unused storage capacity to contain the data from a single disk unit within each pool. If, when the computer system is powered on, there is not enough unused disk storage capacity to hold the data from any single disk unit attached to the system (or within

each logical pool of disk units), then the system operator will receive a warning message, but the system will continue to run normally.

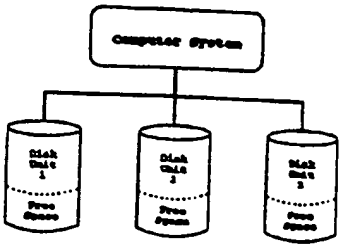

Fig. 3a

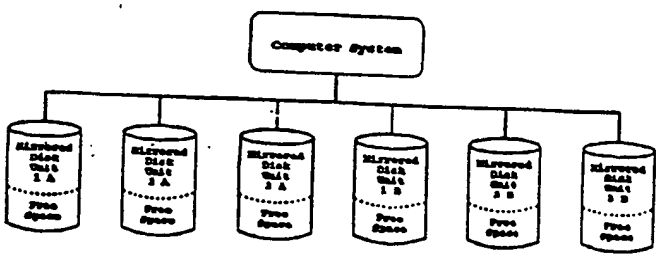

Fig. 3b

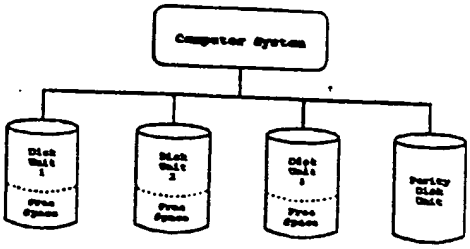

Fig. 3c

Fig. 3 Computer Systems With Virtual Hot Spare Disk Units.

Based on these assumptions, the virtual hot spare disk units approach works as follows:

- For unprotected disk units (Fig. 3a), when a disk unit signals that it is about to fail, the operating system will move all data off that unit and into what was previously unused storage capacity on other active disk units attached to the system (or in the same logical pool of disk units). Then the operating system will remove the disk unit that had predicted its failure from the system configuration.
- For mirrored disk units (Fig. 3b), when one unit fails, the operating system will move the data off the surviving mirrored partner and into what was previously unused storage capacity on other active disk unit mirrored pairs attached to the system (or in the same logical pool of disk unit mirrored pairs). Then the operating system will remove the pair of units that includes the failed unit from the system configuration. This reduces the duration of running exposed to the time it takes to move the data.
- For RAID protected disk units (Fig. 3c), when one unit fails, the operating system will move the data from the failed unit (RAID allows continued read and write operations to the failed unit) and into what was previously unused storage capacity on other active disk units attached to the system (or in the same logical pool of disk units). Then the operating system will -

remove the failed unit from the system configuration and instruct the RAID controller to remove the unit from the RAID array. This may require moving some additional data if the unit being removed from the array contains some of the RAID parity information, as this parity information will now require space on another unit. The duration of running exposed is reduced to the time it takes to move the data and to re-calculate the RAID parity information.

- There may be unique requirements if certain disk units must be attached to particular disk unit controllers. For example, the disk unit which must be read in order to load the first system programs into memory when the computer system is powered on may have to be attached to a particular disk unit controller. In the case of this example:
  - If an unprotected load source disk unit signals that it is about to fail, then the operating system will first move the data off one of the other disk units attached to the particular required controller into what was previously unused storage capacity on other active disk units attached to the system (or in the same logical pool of disk units). It will then move the data from the load source unit onto the just "freed up" unit attached to the particular required controller. Finally, it will remove the original load source unit from the system configuration.
  - If a mirrored load source disk unit fails, then the operating system will first move the data off another mirrored pair of units attached to the particular required controller into what was previously unused storage capacity on other active disk unit mirrored pairs attached to the system (or in the same logical pool of disk unit mirrored pairs). It will then assign one of the just "freed up" units as the new mirrored partner of the surviving load source unit and resynchronize the mirrored data. Finally, it will remove the failed unit and the other "freed up" unit from the system configuration.

This approach to hot spare disk units has the following advantages when compared with the more typical hot spare approach:

- There need be no added expense due to requiring multiple spare units. As long as there is adequate disk capacity attached to the system (or within each logical pool), then virtual hot spare can replace any disk unit. Thus:
  - The hot spare need not be of particular type or size of disk unit.
  - If there is adequate disk capacity, virtual hot spare can replace multiple failed (or about to fail) disk units at once.
- System performance benefits from all attached disk units.
  - Since the disk space that makes up the virtual hot spare is actually unused capacity on active disk units, all disk units attached to the computer system contribute to the performance of the system.